GERMANY/HONG KONG JOINT RESEARCH SCHEME
THE PROJECT REPORT
(for Project Completion)

Project Number: G_HK018/11

Title

Energy-Efficient Data Management with Novel Storage Technologies

Particulars

	Hong Kong team	German team
Name of Project Co-ordinator (with title)	Prof. Jianliang Xu	Prof. Theo Härder
Name of Co-Investigator (if any)	Dr. Haibo Hu	
Institution or Institutional affiliation	[ ] CityU [ ] HKU [ ] CUHK [ ] HKUST [X] HKBU [ ] LU [ ] HKIEd [ ] PolyU	[X] University of Kaiserslautern [ ] Others: _____
Other project team members (if any)		Yi Ou

Funding Period

	1st year	2nd year (if applicable)
Start Date	Jan. 1, 2012	Jan. 1, 2013
Completion Date	Dec. 31, 2012	Dec. 31, 2013

Objective(s) as per original application

1. To explore the storage model design for database systems with the new storage technologies including NAND flash and PCM;
2. To develop energy-efficient indexing and query processing techniques under new storage models;
3. To propose energy-efficient transaction processing methods that leverage fast access and persistency features of PCM;
4. To evaluate the performance of the proposed techniques and methods by combining theoretical analysis and simulation-based experiments.

Details of Report [Please attach relevant document(s)]

i) Outline of proposed research and results obtained

Due to high access performance and price-per-byte considerations, NAND flash memory has been recommended for use as a mid-tier cache in a multi-tier storage system. We identified several weaknesses of the indirect use of flash memory in mid-tier caching. As improvements, we have proposed two new approaches, an indirect approach called LPD (logical page drop) and a native approach called NFA (native flash access). The basic idea is to drop cold pages proactively so that the garbage collection overhead can be minimized. Our experiments demonstrated that both approaches, especially the native one, effectively improve the use of flash memory in the mid-tier cache. NFA reduces the number of garbage collections and block erasures by up to a factor of five and improves the mid-tier throughput by up to 66%.

Phase Change Memory (PCM) is another promising next-generation memory technology with a range of interesting properties: it is non-volatile, bit alterable, and byte addressable. Instead of being used as an external storage solution (like NAND flash memory), PCM is more likely to be used in the main memory system as buffer pools. We have investigated how the database buffer manager can deal with the write endurance problem, which is unique to PCM-based buffer pools. We have designed a family of new buffer algorithms addressing this problem, called wear-aware buffer algorithms, and studied their behaviour using trace-driven simulations.

We have also developed a novel logging scheme named PCMLogging that exploits PCM for both data caching and transaction logging to minimize I/O accesses in disk-based databases. Specifically, PCMLogging caches dirty pages/records in PCM and further maintains an implicit log in the cached updates to support database recovery. By integrating log and cached updates, PCMLogging enables simplified recovery and prolongs PCM lifetime. Furthermore, using PCMLogging, we have designed a wear-leveling algorithm, that evenly distributes the write traffic across the PCM storage space, and a cost-based destaging algorithm that adaptively migrates cached data from PCM to external storage. Compared to classical write-ahead logging (WAL), our trace-driven simulation results revealed up to 1~20X improvement in system throughput.

ii) Significance of research results

With the continuous improvement of CPU clock rates and rapid development of other processor-related technologies (e.g., multi-core, GPU, etc.), the performance gap between processing units and storage devices have been enlarged in recent years. To minimize the gap, an increase of main-memory capacity is required to keep up with the performance growth of processors. At the same time, the explosion of data to be managed also increases the demand on the main memory capacity. As such, there is an urgent need to leverage new storage technologies to optimize the system performance with innovative designs. As demonstrated, our proposed caching and transaction techniques based on NAND flash and PCM storage technologies can substantially reduce the I/O cost and improve the overall system throughput. They require a minimal modification of the existing system architecture and have a good potential to be integrated into real data management systems.

Research output

1. Y. Ou, J. Xu, and T. Härder. "Towards an Efficient Flash-Based Mid-Tier Cache." *Proc. 23rd International Conference on Database and Expert Systems Applications (**DEXA '12**)*, Vienna, Austria, September 2012, pp. 55-70.
2. Y. Ou, L. Chen, J. Xu and T. Haerder. "Wear-Aware Algorithms for PCM-Based Database Buffer Pools." In *Proc. the Second International Workshop on Big Data Management on Emerging Hardware (**HardBD**)*, Macau, June 2014.
3. S. Gao, J. Xu, T. Haerder, B. He, B. Choi, H. Hu. "PCMLogging: Optimizing Transaction Logging and Recovery Performance with PCM", Submitted to *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 2014 (under major revision).

iv) Potential for or impact on further research collaboration

Building on the successful collaboration established during this project, we plan to further investigate PCM-assisted indexing and query processing in hybrid storage models. Since PCM can endure only a limited of writes, a re-design of index structures and related query algorithms has to be considered. Data placement is another issue to be investigated. For example, it would be desirable to use PCM for write-intensive index data and NAND flash for read-intensive application data. We also plan to implement the proposed techniques in an open-source database management system to demonstrate their feasibility. It is of particular interest to incorporate them into SQLite, a database widely used on smartphones, because PCM has already been deployed on mobile handsets.